SEQUENCE LISTING

(1)	GENERAL	INFORMATION
-----	---------	-------------

1	(;)	APPLICANT
		MEEDICANI

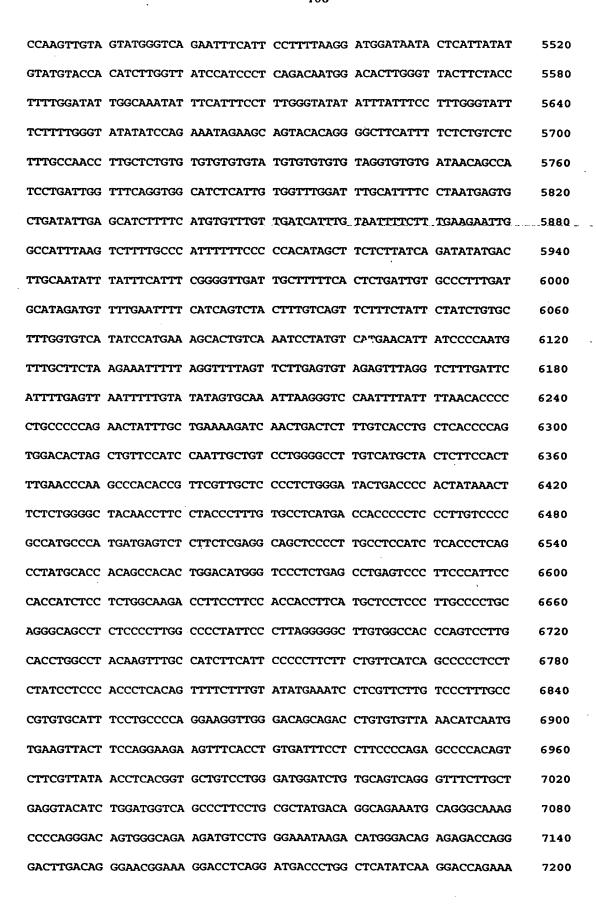
- (A) NAME: Fred Hutchinson Cancer Research Center, Inc.
- (B) STREET: 1100 Fairview Avenue North, Mailstop c2M-027
- (C) CITY: Seattle
- (D) STATE: Washington
- (B) COUNTRY: USA
- (F) POSTAL CODE (ZIP): 98109
- (A) NAME: Thomas Spies
- (B) STREET: 2429 E. Aloha
- (C) CITY: Seattle
- (D) STATE: Washington
- (E) COUNTRY: USA
- (F) POSTAL CODE (ZIP): 98112
- (A) NAME: Veronika Spies
- (B) STREET: 2429 E. Aloha
- (C) CITY: Seattle
- (D) STATE: Washington
- (E) COUNTRY: USA
- (F) POSTAL CODE (ZIP): 98112.
- (ii) TITLE OF INVENTION: CELL STRESS REGULATED HUMAN MHC CLASS I GENE
- (iii) NUMBER OF SEQUENCES: 16
- (iv) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
- (vi) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/029,044
 - (B) FILING DATE: 29-OCT-1996
- (2) INFORMATION FOR SEQ ID NO: 1:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 11722 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CACTGCTTGA	GCCGCTGAGA	GGGTGGCGAC	GTCGGGGCCA	TGGGGCTGGG	CCCGGTCTTC	60
CTGCTTCTGG	CTGGCATCTT	CCCTTTTGCA	CCTCCGGGAG	CTGCTGCTGG	TGAGTGGCGT	120
TCCTGGCGGT	CCTCGGCGGA	GCGGGAGCAG	TGGGACGTTT	CCGGGGGTCG	GGTGGGTAGC	180
GGCGAGCGCT	GTGCGGTCAG	GGCGGGGCTC	СТСТССССТС	TCGGTGGCGC	AGGGAGGTCG	240

ACGCGGCCCG TTACCGCCAC ACTTCAGCCC TGCTTCCCCG TCACTTTTCA GTCCTCCTCG 300 GGATCGCGCA TCACCTGCAC TTTCTGGTCT CCTCCTGCTC TTTCTCTCT CGCGTCTCCT 360 CCGCTTCCTC TCACTTTTCG GACAAACCAG TCCTTCTGAG GCCCATGGGT TCCCGGGCTG 420 CCTCCGGGGC TGCTCCTGTG AATGGCATTC GAGTGCCCTT CCAGCGCGGC CACTGAAGCA 480 GCCACAACCC CCGGTGCTCG GGGCGGCTCT CAGGTCCCTG AAGTCCTGTC CTCTCCCGGA 540 GCCGACGTGT TCTCAGCTCC TGGGCCGCAG CTCCTGGAGT AGGGGCCCTC CTTTCTCGGG 600 ACCCGGAGCT GGTGCTTCCT GCTGCTGTGG GGACTGTGGG GGGTCCTGAC TCTCAAGCTG 660 AGGGGTTGGA GTCTGCAGGC TCCGGGCAGA GGATTCTTCC TGCGACTTCT CTCATCCCCA 720 GCTCATTCTC CCCTCGCCTC TGGCTCCGAG GGTCCTCTCC TCTCTCTCAT CCCACCCCTA 780 CTAATGACCA GTGATCTAAG GACACCAGAT TCCCTCTCAC CTCCTCCCTG CCCATCTCAG 840 GGCCCGCTGA GTCCTTTTGC CCTCCCAGCT CCCTGCTACC CCTTCCTGTG TGCTGTTCTC 900 TGATCCATTT CTAGGGTGTC CTCTGCCCTC ATCCCCTGTC CCCGCCACCG AAGTCCCTCC 960 TGCACCCCTT ATGGGCCTTT CCTACAAGCA GCCTTCACCC AGTGCTGCCC CTATGCCTCC 1020 CCGTTCCCAA ATGTCCCTGA CTCTAACTTT CTGGTGCTGC CTTTTATCCG GGGGGGTCTT 1080 CCCTCCATCC CACTCCCCTC CAGACCCCCA AGGGGAACCC TGATGCTAAT GGCAGTTGGG 1140 CCTTAGGCAG GGCGCAGGGC AGCGCAGATG CCCCCTCCC TCCAGTGCAG ATGCCTGTTC 1200 TGGACCCTGC CTCATTGTGG CCCCTTCCCC ACTCCTTCAT CCTCAGCCTC ACCCTCTTGA 1260 GGACCCCACC CTCCAGCCCA CAGGTGCTGG ACCATCCCTC CCTGGTCCCT CCGCCCCTCT 1320 CCACCTTGGG ACCTTGTGCT GCTCCTATCT CTTGCCCAGC TGCCTTGGGC CCTCAGCACG 1380 TTCTCATCTT TCAGTGGGAA AGTGGGAGTG CTGGAGCATA TGACAGTGCT GAGCATCTTT 1440 CCCAAGCCC ACCCTCCCC AGAGCACCCT CCCCTCCTGT CCTCACCCTA CCCCAAGTTC 1500 TCCCACAGTC ACTCCTGCCC CATGCTCATG CCGCCCTCCA GTTCTTGCTC TGCCCATCTC 1560 CCCTCCCAA CCCAGACCTA AAACAGGCTG TTGGGCCAAC TGTTCCTTGA CCTTCCTTCT 1620 TTTCTTTTGG TTCCTTGACC CCAGTGGGCT CTCACTCCCC ACACCGCATA TCTAAAATCT 1680 GTTTTGCCTG CTCTTGGGGT GCCACTGCTC CCCCTCCAGC ATTACTCCTT TTGGCAGGTC 1740 CTTCCTCAGG CTGAGAATCT CCCCCTCTAC CTTGGTTTTC TCTCTCTGGC CAGCACCCCC 1800 ACTCCTTGCT TTGTTTTAA TTTTTAACTT TTGTTTGGGT ACGTAGTAGA TATATATGTA 1860 TATATTTATG GGGTACATGG GATATTTTGA CACAGGCCTA CAATATGTAA TAATCACATC 1920 AGGGTAAATG GGTTATATCA CAACAAGCAT TTATCCTTTC TTTGTGCTAC AAACAATCCC 1980

ATTATGCTCT TTCAGTTATT TTTAAATGTA CAATAAATTA TTGTTGACTG TACTCACCCT 2040 GCTGTGCTAT CTACTAGATC TTATTCATTC TAATTATATT TTTGTACCCA TTATTAACCA 2100 TCCCTGCTCC CCCACTCCCC ACTACCCTTC TCAGCCTCTG GTAATCATCA TTCTATTGTC 2160 TCTCCCCATG AGGTCCATTG TTTTAAATTT TGGCTGCCAC AAATAAGTGA GAACATGCAA 2220 AGTTTGTCTG TCTGGGCCTG GGGCTTATTT CACTTCACAG GATGACCTCC AGTTCTTTGC 2280 AAATGACACG ATGGCTGAAT AGTTCTCCAC ATACACATGT ACACCACATT TTCTTTATCC 2340 ATGCGTCTGT TGATGGACAC TTAGATTGCT TGCAGATCTT GGCTACTTTG AATAGTGCTG 2400 CAATAAACAT GGAAAAGTAG ATAGCTCTTT AATATACCGA TTTCCTTTCT TTGGAGTATA 2460 TGCCTAACAG TGGGAGTGCT GGAGCATATG ACAGCTCTAT TGTATTTTTA GTTTTTTGGAA 2520 GAACCTCCAC ATTGTTTCCC ATAGTGGTTG TACTAGTTTA CGTTCCCACC AACAGTGTAC 2580 ATCCTCACCA GCATTCCTTA TTTCTACATC CTCGCCAGCA TTCCTTATTG CCTGTCTTCT 2640 GGATAAAAGC CAGTTTATCT GGGGTGGGAT GTTATCTCGT AGGAGTTTTG ATTTGCCTTC 2700 ATCTGTTGAC GAATGATGTT GAGCACCTTT TCATATACCT GTTTGCCATT TATATGTCTT 2760 CTTTTGAGAA ATGACTATTC AGATCTTTTC TCATTTTTAA ATTGGATTAT TATATTTTTT 2820 TTCCTATAGT TGTTCGAGCT CCTTATATGT TTCAGTTACT GATCCTTTGT CAGATGAATA 2880 GTTTGAAAAT ATTTTCTCCC ATTCTTGGAT GGTCTCTTCA TTTTGTTTAT TGTTTCCTTT 2940 GCTGTGCAGA AGCCTTTTTA CTTGATATGA TCCCATTTAT GCAATTTTAC TTTGGTTACC 3000 TGTGCTTGTG GGGTATTACT TTAAAAATCT TTGCCCAGTC CAATATCCTA GAGAGTTTCC 3060 CCAATGTTTT CTTGTATAGT TTCATAGTTT GAGGTCATAG ATTTACATCT TTAATCCACT 3120 TTGATTTGAT TTTTGTATAT GGTGAAAGAC AGGGTCTAGT TTCATTCTTC TGCATAAGGA 3180 TATCTAGTTT CCCCAGCACC ATTTTTGAAG AGACTCTCCT TTGCCAATGT GTGTTCTTGG 3240 TACCTTTGTT GGAAATGAGT TTACTGTAGA TGTATGGAAT TGTTTCTGGG TTCTCTATTC 3300 TGTTTCATTG GTCTGTGTGT CTGTTTTTAT GCCAGTATCA TGCTGTTTTG GTTACTGTAG 3360 CTCTGTAGTA TAATTTGAAG TCAGATAATG TGATTCCTCT AGTTTTGTTC ATTTTGCTCA 3420 GGATAGCTTT ATCTATTCTG GTTTTTTTGT GGTTCCATAT GCATTTTAGG ATTATTTTTA 3480 TTATTTCTGT GAAGAATGTC ATTAGTGTTT TGATAGGGAT TGCATTGAAT CTGTAGATTA 3540 CTTTGGGTAG TATGGATATT TCAACAAAAC TGATTCTTCC AATCCATGAA CGTGGACTAT 3600 CTTTTCCATT TTTTGTGTCC TTCAATTTTT TGCATCAGTG TTTTTTGTTT TTGGTTTTTG 3660 AGATGGAGTT TCACTCTTGT TGCCCAGGCT AGAATGCAAG GGTGTGATCT TGGCTCACCG 3720

CAACCTCCGC CTCCCAGGTT CAAGCTATTC TTCTGCCTCA GCCTCCCAAG TAGCTGGGAT 3780 TACAGGCATG TGCCACTGTG CCTGGCTAAT TTTCTATTTT TATTAGAGAT GGGGTTTCTC 3840 TATGTTGGCC AGGCTAGTCT TGAACTCCTG ACCTCAGGTG ATCCACCTGC CTCGGCCTCC 3900 CAAAGTGCTG GGATTACAGG CATGAGCCAC CACGCCCAGC CACATCACTG TTTTATAGTT 3960 4020 TGTAGCTATT GTAAATGGGA TTCGTTTCTT GATTTCTTT TCAGATTATT TGCTGTTAGC 4080 ACTGATTTIT GCATGTTGAT TTTGTATCCT GCAACTTTAC TGAATTTGTT CTTCAGTTCT 4140 AATGGTTTTT TGGTGGAGTC TTTAGGTTTT TCCAAATATC AGACCACATG ATCTGCAAAC 4200 AAGGATAATT TGACTTCTTC TTTTCCAGTT TTAATGCCCT TTCTTTCTTT CTCCTGTCTG 4260 ATTGCTCTAG TTAGGATCTG CAGTACTGTG TTGCATAACT GTGGTAAAAT TAGTCATCCT 4320 TGTCTTATTC CAGATCTTAG AGAAAAGGCT TTCAGTTTTC CCCCATTCAG TATGTTACTA 4380 GCTGTGAGTT TGTCATATAT GGCTTTTATT ATATTGAGGT CTGTTCCTTG TATACTTAGT 4440 TTTTTGAGAG TTTTTATCAT GAAGGGATGT TGAATTTATC AAATGCTTTT TCAGTATCAA 4500 TTGAATGATA CTGGCTTTTG TCCTTTATTC TGTTGATATG ACGTATTACA TTGATTGATT 4560 TGTGTATGTT AAATCATCCT TGCATACCTG GAATACATTC CACTTGCTCA TAAAGAATGA 4620 TCTTTTTAA TGTATTGTTG AATGTGGTTT GCTAGTATTT CCTTGACGAT TTTTGCATCG 4680 GTGTTCATCA GGGATATAGG CCTGTAGTTT TCTTTTTTAT GATGTGTCTT TGCCTGGTTT 4740 TTGTATCAGG ATATTCCTGG CTTTGTAAAA TGAGTTTGGA AGTATTCCCT CCTCCTCTAT 4800 TTTTCAGAAC AGTTTGAATA GGACTGACAT ATGTTGTTCT TTAAAAGTTT AATTGTGGTA 4860 AATTATACAT TACATAAATT TTACTGTTTT AACCACTTTT AAGTGTATAC TCGGTGGCAT 4920 TAGATACATT CACATTTTTG TGCAACCCAA AACTCTGTGC CCATTAATCG GTAACTCCCC 4980 ATTCCTCCCT ACCTCTGGCC CCTGGTAACC ACCATTCTAC TTTTTGTTTC TATGAATTTG 5040 ACCACTCTAG GTACCTCATT TAAGCAGAAT CATGTAATGT TTGTCTTTTT GTTTCTGGCT 5100 TATTTCACTT ATAATATTTT TGAGGTTCGG TGGGCACAGT GGCTCACGCC TGGATTTCCA 5160 GCACTTTGGG AGGCTGAAGC AGGTGGATCA CCTGAGTTTC GGAGTTCGAA ACCAGCCTGG 5220 CCAACATGGT GAAACCCCAT CTCTACTAAA AATAATAAAA GTTAGCCGGG CGTGATGGCG 5280 GGTGCCTGTA ATCCCAACTA CTTGGGAGGC TGAGGCAGGA GAATCGCTTG AATCCGGGAA 5340 GTGGAGGTTG CAGTGAGCTG AGATCAGGCC ACTGCACTCC AGCCTGGGCA ACAAGAGTGA 5400 AATTCCATCT CCAAAAAAA AAAATAAAAC AATAATAATA ATAATATTT TGAGGTTCAT 5460



GAAGGTGAGA GTCGGCAGGG GCAAGAGTGA CTGGAGAGGC CTTTTCCAGA AAAGTTAGGG 7260 GCAGAGAGCA GGGACCTGTC TCTTCCCACT GGATCTGGCT CAGGCTGGGG GTGAGGAATG 7320 GGGGTCAGTG GAACTCAGCA GGGAGGTGAG CCGGCACTCA GCCCACACAG GGAGGCATGG 7380 GGGAGGCCA GGGAGGCGTA CCCCCTGGGC TGAGTTCCTC ACTTGGGTGG AAAGGTGATG 7440 GGTTCGGGAA TGGAGAAGTC ACTGCTGGGT GGGGGCAGGC TTGCATTCCC TCCAGGAGAT 7500 TAGGGTCTGT GAGATCCATG AAGACAACAG CACCAGGAGC TCCCAGCATT TCTACTACGA 7560 CAGAGCTCAG ACCTTGGCCA TGAACGTCAG GAATTTCTTG AAGGAAGATG CCATGAAGAC 7680 CAAGACACAC TATCACGCTA TGCATGCAGA CTGCCTGCAG GAACTACGGC GATATCTAGA 7740 ATCCAGCGTA GTCCTGAGGA GAAGAGGTAC GGACGCTGGC CAGGGGCTCT CCTCTCCCTC 7800 CAATTCTGCT AGAGTTGCCT CACCTCCAAG ATGTGTCCAG GGAAACCCTC CCTGTGCTAT 7860 GGATGAAGGC ATTTCCTGTT GGCACATCGT GTCCTGATTT TCCTCTATTG TTAGAGCCAC 7920 TGGATAAAGA CAGTGGGTCA GGGACTGGAC CATCCAGTGT TGTAATCAGG GCAAGTAGAG 7980 GACCCTCCGA CAGAATCCTG AGCCTGTGGT GGGTGTCAGG CAGGAGAGGA AGCCTTCAGG 8040 GCCAGGGCTG CCCCCTCTGC CTCCCAGCCT GCCCATCCTG GAGAGTTCCC TCCTGGCCCC 8100 ACAACCCAGG AGTCCACCCC TGACATCCCC CTCCTCAGCA TCAATGTGGG GATCCCAGAG 8160 CCTGAGGCCA CAGTCCCAAG GCCCATCCTC CTGCCAGCCT GGAAGAACTG GGCCCCAGAG 8220 TGAGGACAGA CTTGCAGGTC AGGGGTCCCG GAGGGCTTCA GCCAGAGTGA GAACAGTGAA 8280 GAGAAACAGC CCTGTTCCTC TCCCCTCCTT AGAGGGGAGC AGGGCTTCAC TGGCTCTGCC CTTTCTTCTC CAGTGCCCCC CATGGTGAAT GTCACCCGCA GCGAGGCCTC AGAGGGCAAC 8400 ATCACCGTGA CATGCAGGGC TTCCAGCTTC TATCCCCGGA ATATCACACT GACCTGGCGT 8460 CAGGATGGGG TATCTTTGAG CCACGACACC CAGCAGTGGG GGGATGTCCT GCCTGATGGG 8520 AATGGAACCT ACCAGACCTG GGTGGCCACC AGGATTTGCC AAGGAGAGGA GCAGAGGTTC 8580 ACCTGCTACA TGGAACACAG CGGGAATCAC AGCACTCACC CTGTGCCCTC TGGTGAGCCT 8640 AGGGTGACCC TGGAGAGGGT CAGGCCAGGG TAGGGACAGC AGGGATGGCT GTGGCTCTCT 8700 GCCCAGTGTA TAACAAGTCC CTTTTTTTCA GGGAAAGTGC TGGTGCTTCA GAGTCATTGG 8760 CAGACATTCC ATGTTTCTGC TGTTGCTGCT GCTGCTGCTG CTATTTTTGT TATTATTATT 8820 TTCTATGTCC GTTGTTGTAA GAAGAAAACA TCAGCTGCAG AGGGTCCAGG TGAGAAAAGC 8880 GGGCAGTTTC TGGAGATGGT AAGGCCCCTG TCTGGGCAGT AGGGTCCCCT CATTGCTCCT 8940

GCAAAGATAG GCATGTTGGT GACAAGGCTT CTGTAACAGG GGATGAAAGT TGGGGAATTT 9000 GGGAAGGGAA TGGGGGCAGC ATCTCCATCT ACACCCATAA GTGCTGCCCA AGCGAGGGTC 9060 AAACGCCCAG CTGTGGCATC TTCCTGCTGC AGGTGAGGAG TGGGCAGCAG GGAGGGCTGC 9120 GGCGCCTGCT CTGTCCCCAT CCCGGTCTCT GTGTCTCTTG GACTCACTAG GGCGCATCCA 9180 GGTGGGGTGA GCTGGGAATC ACGTGCTGAA TGCTGAGGGC CTGGATGATC ACGGCCTCAG 9240 AGGGAGCAAA TAGTAAAGGC AGCTGTGATC TGGGGAGGGC CAGAAACTGG AGAGGAATCT 9300 GAGGAGAGGC GGTGCCCCTA TTCCCTTCCT CTCTGCATCC CCCTCCCCTG TTTCTCCAGC 9360 CATCGGGGCG GACACCGAGA AAAAGACCTA TGAGGCCCAG CCTGGGGGCC CTGCCTGTGT 9420 AGCCCTTGG AGACCCCTAG TAACAGGGAG GGTCCTGAGC ACACATGGCC ATCTCTGTCC 9480 ACTGTGCAGC TCCCCATGCA CCTCCTCCAG GAGCTTTCTT GGGGTTGTCG TGTCCTCTGC 9540 ACCATTCGAG GCCCTACTCT TTCCAGGTTC CCACGGCCTG GCCTCCCTGA GTTTCTTGCA 9600 GATGACATGG ATGAGTAGAT AAGCAGATGT CCCTGGGCCA TTTGAGGAGT GGGGCCCAGC 9660 CCCTCATCAG GGCAGCTGTG GTCCCTGTTT TCATCCTACC TCCGAGTGTT TTCTTCTCCA 9720 GTCCCTGAGG GACACAGTCC TCAGGGCCCA TGTTTTTGGG GATTTAATCT GTGCTCTGTG 9780 GCCTCACCTT GCCTTCCCTG AGCCAATTTC CCTTTCTAAA GGTGGTCACT GCCTGGTAAG 9840 TTTGGAGTAA GGGACGGTCA GAATCATTTC CCCTACAGTC AGGTTGTTTG ATGGGGGATG 9900 AAAAGAGACA GCAGGAAGTT TTGTGTTTCT GCAAAGACAG AAGCAGTTCA GGCGACAGTA 9960 AGAGGCTGGG GTGTCCAGGA GGGTGTGTCT GGCAGTAGGG TCGCTGGTTT CTCATCCTTG 10020 AACCTAATTG CACTGTCAGT CGGCCCCTCA GGCCTGAGCA GATGGGAAGG TTTGTCCCCT 10080 GCCCTGCAGC AAGAGGGCCC TGTCCAGGAG GCACCCACAA CAGAGGCAGT GCAGGTCTGT 10140 GGTCACTCCT ACTCTCACCT GTGGCGTCTC CCGTAGAGGG ATTGTCAGTT CTGGTTCCCT 10200 GTGGGCAGGA ATGGTTTCCT CATAGGTCAC TGGAGTTTTG GCCAGGAAAA GAGTATGAAG 10260 TTCATGTGGC AGTTTCTCAA AATTCCTGCT TTCAATGTTG ATGTCCAGTA AAGATATTCG 10320 TAATITCAGC TCTATAATCT TAATAGGATT TCCTCTAATA TTGTGAAGCA TATTATATGA 10380 AACAGGAACA CAAATTCCTC AAAATTCCTG CGATGTCCAA TAAAGATTTT CATAATTTCA 10440 GCTCTGCAAT CTTAATAGGA TTTCCTAATA CTGTAAAGCA TATTAAATGA AACAGGAACT 10500 CAAATTTGGA GCCCCCTCTC CAGGAGGTTC TGTGTGGAGA TGGTGGCTGT GGCAGTGGCA 10560 GTTCCCAGGT GCAGAGGGTG GGCAGAGGCA GCCTCAGGCT AAGGGGTCTC CCCTACTCCA 10620 CATGGAGAAA ATCCCTTGTA GGTTGCAAGG GCAGTGGCCG GGTGGAATCC CTGCTAGGGA 10680

CAGAGCAGGA	AGGCCTCGCA	GCCTCACCAA	GCAGCAGCCC	TGGGGTGGAG	CTGCGTTTCC	10740
AGGGTTAAGC	GGACCAGGCA	GGAGTAGCGG	TTACTCAAGA	GCAGGTCACA	GGCTTGGGTT	10800
GTGAGGGTCA	GGAGAGGCCA	GGCCTCCTCG	AGCAAGGTGG	GGGTCCCAGG	GTCAGGTCAG	10860
GTGCAGATCC	TGTGGCAGCC	ACGTCTTTCC	ATGCTGGGCC	TGCTGGGCCC	CCCAGGCTTC	10920
CTGATGGGGT	CCCCAGTTAG	GAGCTGCCTG	CTCAGGGCTG	GGAGGGGAGG	AGCACTGAGC	10980
TGCAGATAGA	GGGCAGAGCC	CACAGTGGGC	AGGCCTGCC	CTGGTGTGTA	GGTGCCTCTG	11040
CAGGAGAGGA	GGGCCTGGGG.	ACTGAGAGCA	AGGGTCAGGG	.CCTCTCTTTG.	GGGAGGCCTC	-11100
TCACTGTAAC	AGGACTGGTC	AGGCCTGAGA	GGAGGGCACT	GGGTTCCCTC	TTGGGTCTTG	11160
TCCTTTAGTC	TTGGGGCCCT	TTCCCTCCCT	GCACGATGAG	TGGTGGGCAC	AGGGCACGGG	11220
CTGATGTTGA	TGGAGTGATG	GGAGGGAACT	GGCAGGGGCT	GGGAAAAGCA	AGGAGGGAGG	11280
aàgaaaaag	TGGGGGCCTC	ATCTTCCCTC	AGAGAAAGGG	CAAATCTGGT	TTTGGAGCAA	11340
CTGAAGAGAG	AAAAGTCCCC	AGGGAATAAA	CACAACACTG	CACCCAGTGG	AGCATTTACC	11400
CATTTCCCTC	TTTTCTCCAG	AGCTCGTGAG	CCTGCAGGTC	CTGGATCAAC	ACCCAGTTGG	11460
GACGAGTGAC	CACAGGGATG	CCACACAGCT	CGGATTTCAG	CCTCTGATGT	CAGCTCTTGG	11520
GTCCACTGGC	TCCACTGAGG	GCGCCTAGAC	TCTACAGCCA	GGCGGCTGGA	ATTGAATTCC	11580
CTGCCTGGAT	CTCACAAGCA	CTTTCCCTCT	TGGTGCCTCA	GTTTCCTGAC	CTATGAAACA	11640
GAGAAAATAA	AAGCACTTAT	TTATTGTTGT	TGGAGGCTGC	AAAATGTTAG	TAGATATGAG	11700
GCATTTGCAG	CTGTGCCATA	TT				11722

(2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 385 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Gly Leu Gly Pro Val Phe Leu Leu Leu Ala Gly Ile Phe Pro Phe 1 5 10 15

Ala Pro Pro Gly Ala Ala Ala Glu Pro His Ser Leu Arg Tyr Asn Leu 20 25 30

Thr Val Leu Ser Trp Asp Gly Ser Val Gln Ser Gly Phe Leu Ala Glu 35 40 45

- Val His Leu Asp Gly Gln Pro Phe Leu Arg Tyr Asp Arg Gln Lys Cys
 50 55 60
- Arg Ala Lys Pro Gln Gly Gln Trp Ala Glu Asp Val Leu Gly Asn Lys 65 70 75 80
- Thr Trp Asp Arg Glu Thr Arg Asp Leu Thr Gly Asn Gly Lys Asp Leu 85 90 95
- Arg Met Thr Leu Ala His Ile Lys Asp Gln Lys Glu Gly Leu His Ser 100 105 110
- Leu Gln Glu Ile Arg Val Cys Glu Ile His Glu Asp Asn Ser Thr Arg 115 120 125
- Ser Ser Gln His Phe Tyr Tyr Asp Gly Glu Leu Phe Leu Ser Gln Asn 130 135 140
- Val Glu Thr Glu Glu Trp Thr Val Pro Gln Ser Ser Arg Ala Gln Thr 145 150 155 160
 - Leu Ala Met Asn Val Arg Asn Phe Leu Lys Glu Asp Ala Met Lys Thr 165 170 175
- Lys Thr His Tyr His Ala Met His Ala Asp Cys Leu Gln Glu Leu Arg 180 185 190
- Arg Tyr Leu Glu Ser Ser Val Val Leu Arg Arg Arg Val Pro Pro Met 195 200 205
- Val Asn Val Thr Arg Ser Glu Ala Ser Glu Gly Asn Ile Thr Val Thr 210 215 220
- Cys Arg Ala Ser Ser Phe Tyr Pro Arg Asn Ile Thr Leu Thr Trp Arg 225 230 235 240
- Gln Asp Gly Val Ser Leu Ser His Asp Thr Gln Gln Trp Gly Asp Val
- Leu Pro Asp Gly Asn Gly Thr Tyr Gln Thr Trp Val Ala Thr Arg Ile
 260 265 270
- Cys Gln Gly Glu Gln Arg Phe Thr Cys Tyr Met Glu His Ser Gly
 275 280 285
- Asn His Ser Thr His Pro Val Pro Ser Gly Lys Val Leu Val Leu Gln
 290 295 300
- Ser His Trp Gln Thr Phe His Val Ser Ala Val Ala Ala Ala Ala Ala 305 310 315 320
- Ala Ile Phe Val Ile Ile Ile Phe Tyr Val Arg Cys Cys Lys Lys Lys 325 330 335
- Thr Ser Ala Ala Glu Gly Pro Glu Leu Val Ser Leu Gln Val Leu Asp 340 345 350

Gln His Pro Val Gly Thr Ser Asp His Arg Asp Ala Thr Gln Leu Gly 355 360 365

Phe Gln Pro Leu Met Ser Ala Leu Gly Ser Thr Gly Ser Thr Glu Gly 370 375 380

Ala 385

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2380 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GGGCCATGGG	GCTGGGCCGG	GTCCTGCTGT	TTCTGGCCGT	CGCCTTCCCT	TTTGCACCCC	60
CGGCAGCCGC	CGCTGAGCCC	CACAGTCTTC	GTTACAACCT	CATGGTGCTG	TCCCAGGATG	120
AATCTGTGCA	GTCAGGGTTT	CTCGCTGAGG	GACATCTGGA	TGGTCAGCCC	TTCCTGCGCT	180
ATGACAGGCA	GAAACGCAGG	GCAAAGCCCC	AGGGACAGTG	GGCAGAAGAT	GTCCTGGGAG	240
CTAAGACCTG	GGACACAGAG	ACCGAGGACT	TGACAGAGAA	TGGGCAAGAC	CTCAGGAGGA	300
CCCTGACTCA	TATCAAGGAC	CAGAAAGGAG	GCTTGCATTC	CCTCCAGGAG	ATTAGGGTCT	360
GTGAGATCCA	TGAAGACAGC	AGCACCAGGG	GCTCCCGGCA	TTTCTACTAC	GATGGGGAGC	420
TCTTCCTCTC	CCAAAACCTG	GAGACTCAAG	AATCGACAGT	GCCCCAGTCC	TCCAGAGCTC	480
AGACCTTGGC	TATGAACGTC	ACAAATTTCT	GGAAGGAAGA	TGCCATGAAG	ACCAAGACAC	540
ACTATCGCGC	TATGCAGGCA	GACTGCCTGC	AGAAACTACA	GCGATATCTG	AAATCCGGGG	600
TGGCCATCAG	GAGAACAGTG	CCCCCCATGG	TGAATGTCAC	CTGCAGCGAG	GTCTCAGAGG	660
GCAACATCAC	CGTGACATGC	AGGGCTTCCA	GCTTCTATCC	CCGGAATATC	ACACTGACCT	720
GGCGTCAGGA	TGGGGTATCT	TTGAGCCACA	ACACCCAGCA	GTGGGGGGAT	GTCCTGCCTG	780
ATGGGAATGG	AACCTACCAG	ACCTGGGTGG	CCACCAGGAT	TCGCCAAGGA	GAGGAGCAGA	840
GGTTCACCTG	CTACATGGAA	CACAGCGGGA	ATCACGGCAC	TCACCCTGTG	CCCTCTGGGA	900
AGGTGCTGGT	GCTTCAGAGT	CAACGGACAG	ACTTTCCATA	TGTTTCTGCT	GCTATGCCAT	960
GTTTTGTTAT	TATTATTATT	CTCTGTGTCC	CTTGTTGCAA	GAAGAAAACA	TCAGCGGCAG	1020
AGGGTCCAGA	GCTTGTGAGC	CTGCAGGTCC	TGGATCAACA	CCCAGTTGGG	ACAGGAGACC	1080
ACAGGGATGC	AGCACAGCTG	GGATTTCAGC	CTCTGATGTC	AGCTACTGGG	TCCACTGGTT	1140

CC	ACTGAGGG	CGCCTAGACT	CTACAGCCAG	GCGGCCAGGA	TTCAACTCCC	TGCCTGGATC	1200
TC!	ACCAGCAC	TTTCCCTCTG	TTTCCTGACC	TATGAAACAG	AAAATAACAT	CACTTATTTA	1260
TTC	STTGTTGG	ATGCTGCAAA	GTGTTAGTAG	GTATGAGGTG	TTTGCTGCTC	TGCCACGTAG	1320
AG <i>I</i>	AGCCAGCA	AAGGGATCAT	GACCAACTCA	ACATTCCATT	GGAGGCTATA	TGATCAAACA	1380
GC2	AATTGTT	TATCATGAAT	GCAGGATGTG	GGCAAACTCA	CGACTGCTCC	TGCCAACAGA	1440
AGO	STTTGCTG	AGGGCATTCA	CTCCATGGTG	CTCATTGGAG	TTATCTACTG	GGTCATCTAG	1500
AGO	CTATTGT	TTGAGGAATG	CAGTCTTACA	AGCCTACTCT	GGACCCAGCA	GCTGACTCCT	1560
TC?	TCCACCC	CTCTTCTTGC	TATCTCCTAT	ACCAATAAAT	ACGAAGGGCT	GTGGAAGATC	1620
AG <i>I</i>	AGCCCTTG	TTCACGAGAA	GCAAGAAGCC	CCCTGACCCC	TTGTTCCAAA	TATACTCTTT	1680
TGT	CTTTCTC	TTTATTCCCA	CGTTCGCCCT	TTGTTCAGTC	CAATACAGGG	TTGTGGGGCC	1740
CTI	TAACAGTG	CCATATTAAT	TGGTATCATT	ATTTCTGTTG	TTTTTGTTTT	TGTTTTTGTT	1800
TT	GTTTTTG	AGACAGAGTC	TCACTCGTCA	CCCAGGCTGC	AGTTCACTGG	TGTGATCTCA	1860
GC7	CACTGCA	ACCTCTGCCT	CCCAGGTTCA	AGCACTTCTC	GTACCTCAGA	CTCCCGATAG	1920
CTC	GGATTAC	AGACAGGCAC	CACCACACCC	AGCTAATTTT	TGTATTTTT	GTAGAGACGG	1980
GG1	TTCGCCA	AGTTGACCAG	CCCAGTTTCA	AACTCCTGAC	CTCAGGTGAT	CTGCCTGCCT	2040
TGG	CATCCCA	AAGTGCTGGG	ATTACAAGAA	TGAGCCACCG	TGCCTGGCCT	ATTTTATTAT	2100
AT'I	GTAATAT	ATTTTATTAT	ATTAGCCACC	ATGCCTGTCC	TATTTTCTTA	TGTTTTAATA	2160
TAT	TTTAATA	TATTACATGT	GCAGTAATTA	GATTATCATG	GGTGAACTTT	ATGAGTGAGT	2220
ATC	TTGGTGA	TGACTCCTCC	TGACCAGCCC	AGGACCAGCT	TTCTTGTCAC	CTTGAGGTCC	2280
CCI	rcgccccg	TCACACCGTT	ATCGATTACT	CTGTGTCTAC	TATTATGTGT	GCATAATTTA	2340
TAC	CCGTAAAT	GTTTACTCTT	TAAATAAÁAA	ааааааааа			2380

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 383 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Gly Leu Gly Arg Val Leu Leu Phe Leu Ala Val Ala Phe Pro Phe 1 5 10 15

Ala Pro Pro Ala Ala Ala Ala Glu Pro His Ser Leu Arg Tyr Asn Leu 20 25 30

Met Val Leu Ser Gln Asp Glu Ser Val Gln Ser Gly Phe Leu Ala Glu 35 40 45

Gly His Leu Asp Gly Gln Pro Phe Leu Arg Tyr Asp Arg Gln Lys Arg
50 55 60

Arg Ala Lys Pro Gln Gly Gln Trp Ala Glu Asp Val Leu Gly Ala Lys 65 70 75 80

Thr Trp Asp Thr Glu Thr Glu Asp Leu Thr Glu Asn Gly Gln Asp Leu 85 90 95

Arg Arg Thr Leu Thr His Ile Lys Asp Gln Lys Gly Gly Leu His Ser 100 105 110

Leu Gln Glu Ile Arg Val Cys Glu Ile His Glu Asp Ser Ser Thr Arg 115 120 125

Gly Ser Arg His Phe Tyr Tyr Asp Gly Glu Leu Phe Leu Ser Gln Asn 130 135 140

Leu Glu Thr Gln Glu Ser Thr Val Pro Gln Ser Ser Arg Ala Gln Thr 145 150 155 160

Leu Ala Met Asn Val Thr Asn Phe Trp Lys Glu Asp Ala Met Lys Thr
165 170 175

Lys Thr His Tyr Arg Ala Met Gln Ala Asp Cys Leu Gln Lys Leu Gln 180 185 190

Arg Tyr Leu Lys Ser Gly Val Ala Ile Arg Arg Thr Val Pro Pro Met 195 200 205

Val Asn Val Thr Cys Ser Glu Val Ser Glu Gly Asn Ile Thr Val Thr 210 215 220

Cys Arg Ala Ser Ser Phe Tyr Pro Arg Asn Ile Thr Leu Thr Trp Arg 225 230 235 240

Gln Asp Gly Val Ser Leu Ser His Asn Thr Gln Gln Trp Gly Asp Val
245 250 255

Leu Pro Asp Gly Asn Gly Thr Tyr Gln Thr Trp Val Ala Thr Arg Ile 260 265 270

Arg Gln Gly Glu Glu Gln Arg Phe Thr Cys Tyr Met Glu His Ser Gly 275 280 285

Asn His Gly Thr His Pro Val Pro Ser Gly Lys Val Leu Val Leu Gln 290 295 300

Ser Gln Arg Thr Asp Phe Pro Tyr Val Ser Ala Ala Met Pro Cys Phe 305 310 315 320

Val Ile Ile Ile Leu Cys Val Pro Cys Cys Lys Lys Lys Thr Ser 325 330 335

	Ala	Ala	Glu	Gly 340	Pro	Glu	Leu	Val	Ser 345	Leu	Gln	Val	Leu	Asp 350	Gln	His	
	Pro	Val	Gly 355	Thr	Gly	Asp	His	Arg 360	Asp	Ala	Ala	Gln	Leu 365	Gly	Phe	Gln	
	Pro	Leu 370	Met	Ser	Ala	Thr	Gly 375	Ser	Thr	Gly	Ser	Thr 380	Glu	Gly	Ala		
(2)	INFO	RMAT:	ION 1	FOR 8	SEQ :	ED NO	D: 5:	:									
	(i)	(A) (B) (C)	LEN TYI	NGTH PE: 1 RANDI	: 25 nucle EDNES	reris base eic a SS: s linea	e pai acid singl	irs									
	(xi)	SEQU	JENCI	B DES	SCRII	OITS	1: SI	EQ II	NO:	: 5:							
ACTO	eggga <i>i</i>	AC A	AGGTT	TAT?	A TGZ	AGA											25
(2)	INFOR	TAMS	ON E	FOR S	SEQ 1	D NC): 6:	:									
	(i)	(A) (B) (C)	LEN TYP STR	NGTH: PE: r RANDE	: 24 nucle EDNES	base base ic a S: s	pai cid ingl	.rs					٠				
	(xi)	SEQU	JENCE	DES	CRIE	TION	: SE	Q II	NO:	6:							
TGT	CACCCC	er C	TCT	ACAGO	acc	:c											24
(2)	INFO	RMATI	ION E	FOR S	SEQ 1	D NC): 7:									,	
	(i)	(A) (B) (C)	LEN TYI STI	NGTH: PE: r RANDI	: 17 nucle EDNES	TERIS base ic a SS: s linea	e pai cid singl	rs							•		
	(xi)	SEQ	JENCI	E DES	SCRII	OIT	1: SI	EQ II	ONO:	: 7:							
GGG	GCCAT	GG GG	GCTG	G G	•												17
(2)	INFO	RMAT	ION 1	FOR :	SEQ :	ID NO): 8	:									•
	(i)	(A (B (C) LE) TY) ST	NGTH PE: 1 RAND	: 17 nucle EDNE	reris base eic a SS: s	e pa: acid sing:	irs									
	(xi)	SEQ	UENC:	E DE	SCRI	PTIO	N: S	EQ I	D NO	: 8:							

ATCTGAGATG TCGGTCC	17
(2) INFORMATION FOR SEQ ID NO: 9:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 25 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: single	
(D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:	
CGTTCTTGTC CCTTTGCCCG TGTGC	25
(2) INFORMATION FOR SEQ ID NO: 10:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 25 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: single	
(D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:	
AACCCTTCCC TTACCCCCGT CGTAG	25
(2) INFORMATION FOR SEQ ID NO: 11:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 45 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: single -	
(D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:	
TATGTAAAAC GACGGCCAGT TTCACCTGTG ATTTCCTCTT CCCCA	.45
(2) INFORMATION FOR SEQ ID NO: 12:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 45 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: single	
(D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:	
GGTCTTTCA ATCCCCGTCT CTCGTCCAGT ATCGACAAAG GACAT	45
(2) INFORMATION FOR SEQ ID NO: 13:	

(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 40 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:	
TATGTAAAAC GACGGCCAGT TTCGGGAATG GAGAAGTCAC	40
(2) INFORMATION FOR SEQ ID NO: 14:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:	•
CGAGAGGAGA GGGAGGTTAA CCAGTATCGA CAAAGGACAT	40
(2) INFORMATION FOR SEQ ID NO: 15:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:	
TATGTAAAAC GACGGCCAGT GTTCCTCTCC CCTCCTTAGA	40
(2) INFORMATION FOR SEQ ID NO: 16:	
 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 41 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:	•
AAAAAGTCCC TTTCACGACC ACCAGTATCG ACAAAGGACA T	41